

An Application of Principal Component Analysis on Multivariate Time- Stationary Spatio- Temporal Data

Stephan Stahlschmidt*
Wolfgang Karl Härdle*
Helmut Thome**



* Humboldt-Universität zu Berlin, Germany

** Martin-Luther-Universität Halle-Wittenberg, Germany

This research was supported by the Deutsche
Forschungsgemeinschaft through the SFB 649 "Economic Risk".

<http://sfb649.wiwi.hu-berlin.de>
ISSN 1860-5664

SFB 649, Humboldt-Universität zu Berlin
Spandauer Straße 1, D-10178 Berlin



An Application of Principal Component Analysis on Multivariate Time–Stationary Spatio–Temporal Data*

Stephan Stahlschmidt

School of Business and Economics

Humboldt–Universität zu Berlin

Germany

Wolfgang K. Härdle

C.A.S.E Center for Applied Statistics and Economics

Humboldt–Universität zu Berlin

Germany

Helmut Thome

Institute of Sociology

Martin-Luther-Universität Halle-Wittenberg

Germany

*The financial support from the Deutsche Forschungsgemeinschaft via SFB 649 “Ökonomisches Risiko”, Humboldt–Universität zu Berlin, and IRTG 1792 “High Dimensional Non Stationary Time Series”, Humboldt–Universität zu Berlin, is gratefully acknowledged.

Abstract

Principal component analysis denotes a popular algorithmic technique to dimension reduction and factor extraction. Spatial variants have been proposed to account for the particularities of spatial data, namely spatial heterogeneity and spatial autocorrelation, and we present a novel approach which transfers principal component analysis into the spatio-temporal realm.

Our approach, named stPCA, allows for dimension reduction in the attribute space while striving to preserve much of the data's variance and maintaining the data's original structure in the spatio-temporal domain. Additionally to spatial autocorrelation stPCA exploits any serial correlation present in the data and consequently takes advantage of all particular features of spatial-temporal data.

A simulation study underlines the superior performance of stPCA if compared to the original PCA or its spatial variants and an application on indicators of economic deprivation and urbanism demonstrates its suitability for practical use.

Keywords: dimension reduction, economic deprivation, factor extraction, PCA, spatio-temporal analysis, urbanism

JEL codes: C31, C33, R11

1 Introduction

Factor extraction refers to the process of concentrating several variables into a set of factors with lower cardinality and has been applied in virtually any field of statistical analysis. It denotes a dimension reduction technique, as well as a vehicle to disclose latent factors. Due to the reduction factor extraction relieves the computational burden in any subsequent analysis, might help to avoid the curse of dimensionality and most importantly presents measurements of theoretical interest which would otherwise remain hidden due to incomplete knowledge on the subject matter or due to the latent nature of the variable of interest. Consequently, factor extraction might be understood as an analysis tool, which helps to identify the relevant factors of interest.

Principal component analysis (Pearson, 1901; Hotelling, 1933), which is also known as discrete Karhunen–Loève transformation (Karhunen, 1947; Loève, 1948), Hotelling transformation (Hotelling, 1933) or the method of empirical orthogonal functions (Lorenz, 1956) among others, is frequently applied to extract factors from a set of variables (e.g. Jolliffe, 2002, chap. 4). It is in fact based on a transformation of the data, in which the orthogonal coordinates are rotated in order to load as much variance as possible on the first components and less and less variance on subsequent components. Consequently the first components, formed by a linear combination of the original variables, represent an essential information content of the data and might be understood as factors. By contrast the final components, presenting little residual variance, might be ignored in the analysis and allow thus for dimension reduction. In a strict implementation without any additional rotation and based on standardized variables, principal component analysis (PCA) resembles more an algorithm than a model and restricts the researcher’s influence on choosing the appropriate number of latent factors. This feature distinguishes PCA from other factor extraction techniques, most notably the model-based factor analysis (Spearman, 1904).

However, the application of the PCA algorithm is not exclusively restricted to the attribute subspace, but in case of spatio-temporal data might also be used on the geographical or temporal subspace and consequently reduce either the geographical or the temporal dimension. Richman (1986) proposes a classification of PCA for spatio-temporal data into six modes, where each mode describes exclusive combinations of two subspaces. E.g., the application of PCA on multivariate spatial entities is labeled R-mode and several spatial PCA variants have been proposed (Wartenberg, 1985; Thioulouse et al., 1995; Fotheringham et al., 2002; Jombart et al., 2008). Contrary to the original PCA, these techniques incorporate either spatial autocorrelation or spatial heterogeneity into the PCA approach to factor extraction and the authors demonstrate the superior performance of these spatial PCA variants to disclose any spatial factor, if compared to the original PCA.

On the other hand, these spatial PCA variants only address spatial cross-sectional data and do not apply to spatio-temporal data. In order to allow for a truly spatio-temporal analysis we propose a novel PCA approach, that not only accounts for the spatial peculiarities, but also incorporates serial correlation over time. This spatio-temporal PCA variant (henceforth stPCA) allows for dimension reduction on the attribute space, while preserving the geographical and temporal space, that is, it extracts spatio-temporal factors from several spatio-temporal variables while maintaining the geographical and temporal structure of the original variables.

In the framework of Richman (1986) stPCA can be understood as the combined PR-mode of PCA on spatio-temporal data and the technique describes a transfer of the original PCA to the spatio-temporal realm of geographical and serial correlation. Consequently the proposed technique shares some features with the three-mode PCA of Kroonenberg and de Leeuw (1980), which however relies on i.i.d. observations and has not been studied for correlated observations. Furthermore three-mode PCA includes a dimension reduction in every subspace, whereas stPCA focuses exclusively on the attribute subspace.

The inclusion of latent factors in models for spatio-temporal data is also facilitated by Bayesian hierarchical models (Gelman and Hill, 2006). Recent examples include Tzala and Best (2007); Lawson et al. (2008) and Choi et al. (2012) in public health studies and Hogan and Tchernis (2004) in economics. These models rely on latent factors to regress some explanatory variables on a dependent variable and the latent factors consequently serve as an intermediate step and are not of particular interest in the respective analysis. stPCA consequently represents a novel attempt to incorporate spatial and temporal correlation into a PCA framework and hence facilitates the inclusion of latent factors into spatio-temporal models.

In order to illustrate the performance of stPCA, we present a simulation study and apply stPCA to a data set of economic deprivation and urbanism indicators in Germany. In the Monte Carlo simulation stPCA improves the ordinary and spatial PCA approaches, if a non-negligible spatial structure is present in the spatio-temporal data. The reported difference is substantial and significant. A large gain is made on small n , high t samples, whereas the additional value for large n data seems less pronounced.

The application of stPCA on the indicators of economic deprivation and urbanism illustrates the additional value of a combined spatio-temporal approach if compared to a cross-sectional spatial approach. Only stPCA allows for time specific projections, which highlight the west-east and internal north-south divide in economic deprivation and reliably indicates the big German metropolitan areas.

The following section 2 presents the proposed stPCA approach, which is afterwards evaluated via a simulation in section 3. An actual implementation of stPCA is presented in section 4 and section 5 concludes with a discussion.

2 The spatio-temporal PCA

The original principal component analysis of Pearson (1901) and Hotelling (1933) describes a rotation of the p -dimensional coordinate system. The rotated coordinates present the best orthogonal fit of the data, in which the first coordinate is aligned in the direction of the data's maximum variance. Any subsequent coordinate is afterwards orientated to contain as much of the residual variance as possible conditioned on being orthogonal to all former coordinates.

In this new coordinate system, the coordinates possessing much variance contain most of the data's information, whereas coordinates with a relative small amount of variance contribute little additional information and consequently can be ignored at little cost. This advantage of PCA is facilitated by the orthogonal rotation and allows for dimension reduction in multivariate data while preserving the general

structure of the individual data points.

Upon obtaining the new coordinates the p -dimensional and centred random variables $\mathbf{X} \in \mathbb{R}^p$ are projected onto this new coordinates system via a linear combination. The projections ϕ onto the first coordinate are obtained via

$$\phi = \mathbf{X}u,$$

where u denotes a weight vector which can be identified via the aforementioned variance characteristics of the rotated coordinates. In detail, PCA maximizes the variance in the rotated coordinates, that is, the variance of the projected data points ϕ :

$$\begin{aligned} \max_u \text{Var}(\phi) &= \max_u \text{Var}(\mathbf{X}u) \\ &= \max_u u^\top n^{-1} \mathbf{X}^\top \mathbf{X} u \\ &= \max_u u^\top \Sigma u \end{aligned} \quad (1)$$

where Σ denotes the covariance matrix of the centred \mathbf{X} and the maximization is subject to some identification restriction, like $\|u\| = 1$.

An eigendecomposition of Σ resolves the maximization requirement (1), as the eigenvector corresponding to the largest eigenvalue constitutes the optimal u (Härdle and Simar, 2012). Likewise the projection onto subordinate components is conducted via the remaining eigenvectors, where the corresponding eigenvalues describe the variance explained by this component, and consequently its rank.

The just described original principal component analysis does not address the particularities of spatial data, like spatial autocorrelation or spatial heterogeneity. Spatial extensions to PCA have been proposed, which explicitly account for either heterogeneity (Fotheringham et al., 2002) or autocorrelation (Wartenberg, 1985; Thioulouse et al., 1995; Jombart et al., 2008). In this paper we concentrate on the second type, but would like to note that the suggested spatio-temporal approach might also be adapted to the spatial heterogeneity case.

The suggested extensions amplify the maximization criterion by incorporating the spatial autocorrelation of the projected data points ϕ . Consequently, the proposed methods seek to project the observations onto a new coordinate system, while preserving the spatial relation between the observations and this second objective differentiates the spatial approaches from the ordinary PCA.

Moran's I describes a frequently used statistic of spatial autocorrelation (Moran, 1950) which defines the spatial autocorrelation for some random variable X with mean \bar{X} as

$$I(X) = \frac{N}{\sum_{i=1}^N \sum_{j=1}^N w_{ij}} \frac{\sum_{i=1}^N \sum_{j=1}^N w_{ij} (X_i - \bar{X}) (X_j - \bar{X})}{\sum_{i=1}^N (X_i - \bar{X})^2}, \quad (2)$$

where $w_{i,j}$, drawn from a spatial weight matrix W , describes the spatial weight imposed by observation j onto observation i . Choosing an appropriate spatial weight matrix for either point or areal data and a suitable standardization is up to the subject-matter researcher and might simplify the computation (2).

Indeed the differences of the above mentioned spatial PCA approaches can be attributed to the particular spatial weight matrix and specific transformation of the original variables \mathbf{X} chosen by the authors. In detail, MSC (Wartenberg, 1985) standardizes the original variables and the distance based spatial weight matrix, whereas the Global Structure by Thioulouse et al. (1995) relies on a standardized binary connection matrix and transforms the original data by a mean which is based on assigning weights according to the number of individual neighbours. Finally sPCA (Jombart et al., 2008) applies a row standardization on the binary connection matrix and, because of its specific application to alleles does not standardize the data, but subtracts only the mean.

All these spatial extensions to PCA seek to maximize the product of the variance and spatial autocorrelation of ϕ :

$$\begin{aligned}\max_v \text{Var}(\phi) \text{I}(\phi) &= \max_v \text{Var}(\mathbf{X}v) \text{I}(\mathbf{X}v) \\ &= \max_v v^\top n^{-1} \mathbf{X}^\top W \mathbf{X} v \\ &= \max_v v^\top \Omega v\end{aligned}$$

where $\Omega = n^{-1} \mathbf{X}^\top W \mathbf{X}$ denotes a spatial correlation matrix and the optimal v are found via an eigendecomposition of Ω . Wartenberg (1985) and Thioulouse et al. (1995) point out, that Ω might not be positive definite and state that the resulting negative eigenvalues represent local structure. In case of a non-symmetrical spatial weight matrix W , Jombart et al. (2008) observe, that the optimal v is given by the eigenvector corresponding to the largest eigenvalue of $(2n)^{-1} \mathbf{X}^\top (W + W^\top) \mathbf{X}$.

Spatio-temporal data add another subspace to the attribute and geographical space of spatial data and present measurements of the same multivariate spatial entities over time. Consequently PCA or any spatial PCA variant could be applied at every t , and T eigendecompositions of the time dependent (spatial) covariance matrix could be computed. Hence any serial correlation over time would be ignored and at every t we would have a separate cross-sectional analysis.

Contrarily, stPCA forms a truly spatio-temporal technique. Instead of conducting an analysis at every t separately stPCA proposes to calculate a time average of the spatial covariance matrix and apply an eigendecomposition on this average. Consequently stPCA exploits any serial correlation and makes use of the fact that the repeated measurements on the time stable spatial entities represent the same information content, whereas any additional noise might vary over time. Hence the time-averaged spatial covariance matrix will include a higher signal to noise ratio and present time stable eigenvectors.

This feature of stPCA, contrary to the repeated application of PCA or its spatial variants, will result in consistent signs and order of the components across t and consequently facilitates the interpretation and further use of the findings. Finally, stPCA is faster, as a function of t , than any repeated application of its non-temporal siblings, as the time-consuming eigendecomposition has to be conducted only once instead of t times.

In detail stPCA maximizes the time average of the product between the variance

and spatial autocorrelation of the projected data points ϕ :

$$\begin{aligned}
\max_{\mu} T^{-1} \sum_{t=1}^T \text{Var}(\phi_t) \text{I}(\phi_t) &= \max_{\mu} T^{-1} \sum_{t=1}^T \text{Var}(\mathbf{X}_t \mu) \text{I}(\mathbf{X}_t \mu) \\
&= \max_{\mu} \mu^{\top} \left(T^{-1} n^{-1} \sum_{t=1}^T \mathbf{X}_t^{\top} W \mathbf{X}_t \right) \mu \\
&= \max_{\mu} \mu^{\top} \Theta \mu,
\end{aligned} \tag{3}$$

where $\Theta = T^{-1} n^{-1} \sum_{t=1}^T \mathbf{X}_t^{\top} W \mathbf{X}_t$ denotes a time average of the spatial correlation matrix. If W is symmetric, the optimal weight vector μ may be extracted as before from a direct eigendecomposition of Θ . Otherwise, and along the reasoning of Jombart et al. (2008), the optimal μ may be found by the eigendecomposition of $(2Tn)^{-1} \sum_{t=1}^T \mathbf{X}_t^{\top} (W + W^{\top}) \mathbf{X}_t$. As in the ordinary PCA and its spatial variants the projections ϕ_t of stPCA are obtained via multiplying the original data \mathbf{X}_t with the time stable principal eigenvectors μ .

3 Simulation

We present two simulations which compare the performance of the original PCA, its spatial variants and the novel stPCA approach to detect spatio-temporal factors. In a first step, we apply the distinct PCA variants to an artificial data set of a single, hidden and stable spatio-temporal factor, which is observed via three noisy variables and is obscured by three additional random noise variables. As in Wartenberg (1985) a ratio between the first eigenvalue and the sum of the absolute value of all eigenvalues is presented to reveal the sensitivity of these techniques in detecting the spatio-temporal factor. Obviously, a high ratio indicates that the respective PCA procedure correctly identifies the single predetermined factor present in the simulated data.

In a second simulation we apply PCA, its spatial variants and stPCA to a data set with two different spatio-temporal factors in order to learn the accuracy of these principal component approaches. Each factor exhibits a distinct and stable spatial-temporal pattern, which affects three noisy variables each and is furthermore masked by six additional random noise variables. We check whether the PCA variants identify the correct number of factors, how the PCA variants weigh the original variables in the computation of the projections and compare to what extent the diverse projections match the original factors.

We start the first simulation by generating a spatial structure $\mathbf{S}^{(1)}$ which takes the form of a square grid of size $\sqrt{n} \times \sqrt{n}$. The n observations $S_{i \in \{1, \dots, n\}}^{(1)}$ follow a normal distribution with a mean depending on the grid's column index c :

$$\begin{aligned}
S_{i,c}^{(1)} &\sim \text{N}(0, 1), \text{ for } c \leq \frac{C}{2} \\
S_{i,c}^{(1)} &\sim \text{N}(\delta, 1), \text{ for } c > \frac{C}{2},
\end{aligned}$$

where δ defines an increment and $C = \sqrt{n}$ denotes the number of columns. Consequently we simulate a patch, which differentiates between the left and right side of the grid by the expectation $\text{E} \left[S_{i,c}^{(1)} \mid c \leq \frac{C}{2} \right] = 0$ and $\text{E} \left[S_{i,c}^{(1)} \mid c > \frac{C}{2} \right] = \delta$. This spatial structure is standardized and subsequently introduced as a constant in the

AR(1) process of the spatio-temporal factor $F_{i,t}^{(1)}$. Switching to vector notation, the factor $\mathbf{F}_t^{(1)} = \text{vec}(F_{1,t}^{(1)}, \dots, F_{n,t}^{(1)})$ is generated via

$$\mathbf{F}_t^{(1)} = \mathbf{S}^{(1)} + 0.5\mathbf{F}_{t-1}^{(1)} + \boldsymbol{\varepsilon}_t, \quad (4)$$

where we define the error vector by $\boldsymbol{\varepsilon}_t \sim \mathcal{N}(\mathbf{0}_n, 0.75\mathbf{I}_n)$. This simulated factor produces $n \times t$ observations, which exhibit a stable spatial pattern over time defined by the size of the increment δ . A high value of δ will result in a more pronounced spatial pattern and, due to the standardization, will not automatically increase the factor's variance, which is instead defined by the coefficient and error vector in the AR(1) process (4).

In our simulation the standardized spatial factor affects $p_1 = 3$ dependent variables $\mathbf{X}_{t,p_1 \in \{1, \dots, 3\}}$, which are defined by the sum of the factor $\mathbf{F}_t^{(1)}$ and an individual AR(1) noise process \mathbf{Z}_{t,p_1} :

$$\mathbf{X}_{t,p_1} = \mathbf{F}_t^{(1)} + \mathbf{Z}_{t,p_1}.$$

The noise process \mathbf{Z}_{t,p_1} differentiates the three variables \mathbf{X}_{t,p_1} via its error component:

$$\mathbf{Z}_{t,p_1} = 0.5\mathbf{Z}_{t-1,p_1} + \boldsymbol{\varepsilon}_{\mathbf{Z},t},$$

where $\boldsymbol{\varepsilon}_{\mathbf{Z},t} \sim \mathcal{N}(\mathbf{0}_n, 0.75\mathbf{I}_n)$ denotes white noise. Consequently we separate the variables by their specific errors drawn from the same normal distribution

Apart from the dependent variables \mathbf{X}_{t,p_1} , we also add $p_2 = 3$ random noise variables $\mathbf{X}_{t,p_2 \in \{4, \dots, 6\}}$, which are independent of the factors and follow an ordinary AR(1) process:

$$\mathbf{X}_{t,p_2} = 0.5\mathbf{X}_{t-1,p_2} + \boldsymbol{\varepsilon}_{\mathbf{X},t},$$

where $\boldsymbol{\varepsilon}_{\mathbf{X},t} \sim \mathcal{N}(\mathbf{0}_n, 0.75\mathbf{I}_n)$ denotes white noise. Hence these three variables possess the same mean and variance as the spatio-temporal factor and interfere with its disclosure.

We run this simulation in two settings to cover small and large n applications. At first we set $n_1 = 49$ and $t_1 \in \{5, 50, 100\}$. This specification allows for all possible combinations of n and t : $t < n$, $t \approx n$ and $t > n$. The same holds for the second setting, where $n_2 = 400$ and $t_2 \in \{40, 400, 800\}$. The simulation is based on a row weighted binary spatial weight matrix indicating direct neighbours and in order to observe the effect of the spatial increment δ , we increase it gradually via steps of 0.4 in the interval $[0, 8]$ for all combinations of n and t . At $\delta = 0$ obviously no spatial factor is produced, as this particular parametrization describes an i.i.d. scenario. Finally we run each combination of the parameters 1000 times and present the respective mean ratio between the first eigenvalue and the sum of all eigenvalues in Figure 1.

We observe, that PCA present a constant ratio between the largest eigenvalue and the sum of all eigenvalues. This ratio remains unaffected by an increase in the spatial increment δ and hence PCA fails to clearly identify the increasingly pronounced spatio-temporal factor. On the other hand the spatial PCA variants gain strongly from an increase in δ . The initial ratio at $\delta = 0$ is increased more than twofold at $\delta = 8$ and the spatial PCA variants cause higher ratios than the original PCA for $\delta \geq 1.2$ ($n_1 = 49$), respectively $\delta \geq 0.4$ ($n_2 = 400$). Consequently we can verify the

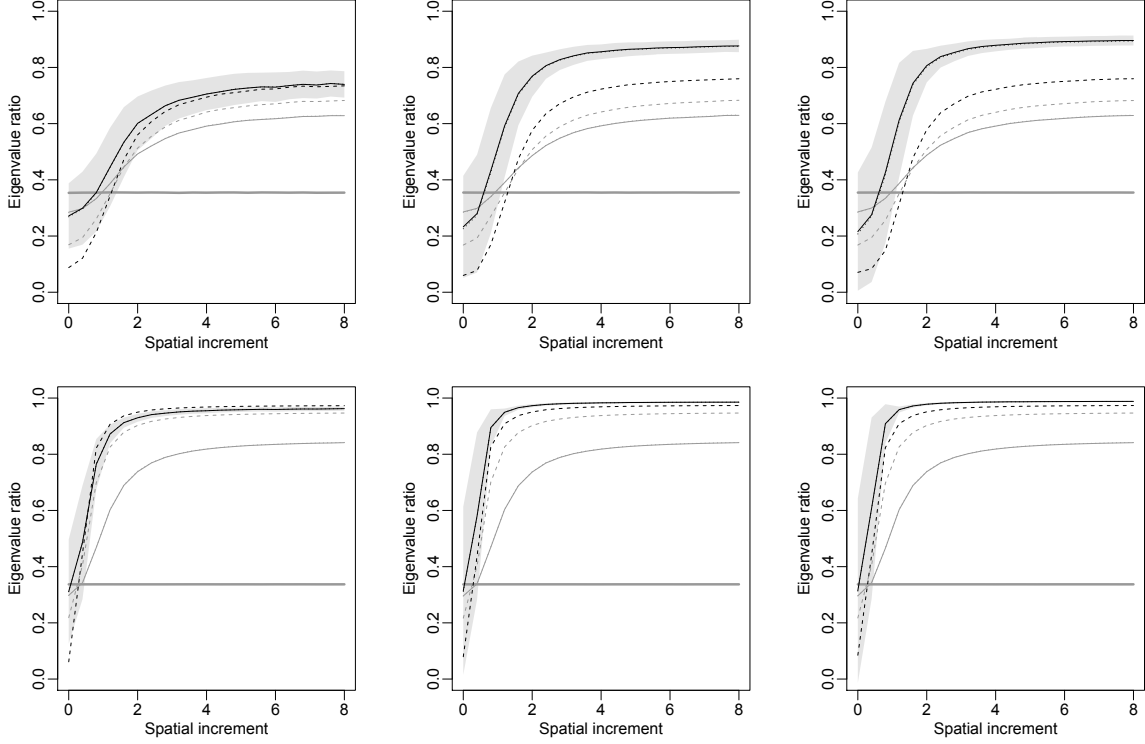


Figure 1: Mean ratio (with standard deviation for temporal sPCA) of the first eigenvalue to the sum of all eigenvalues as assigned by PCA (thick solid grey line), sPCA (solid grey line), MSC (dashed grey line), Global Structure PCA (dotted grey line), temporal sPCA (solid black line), temporal MSC (dashed black line) and temporal Global Structure PCA (dotted black line) for $n_1 = 49$ (first row) with $t_1 = 5$ (left graph), $t_1 = 50$ (middle graph) and $t_1 = 100$ (right graph), and $n_2 = 400$ (second row) with $t_2 = 40$ (left graph), $t_2 = 400$ (middle graph) and $t_2 = 800$ (right graph).

results of Wartenberg (1985), Thioulouse et al. (1995) and Jombart et al. (2008), and observe that extending PCA by a spatial component improves the sensibility of the spatial PCA variants to identify a spatial factor.

However, as can be observed in Figure 1, stPCA is even more responsive to an increase in δ than the spatial PCA Variants. At $\delta \geq 0.8$ ($n_1 = 49$), respectively $\delta \geq 0.4$ ($n_2 = 400$) the ratio reported by stPCA is larger than the ratio of any other PCA variant including the purely spatial PCA variants. In detail, stPCA not only reports larger ratios for any given n and t than the spatial PCA variants at the aforementioned levels of δ , but also exploits an increase in n much stronger. Furthermore only stPCA makes use of the time dimension and reports higher ratios for an increase in t . Unsurprisingly the solely spatial PCA variants do not gain on such an increase in t , but only on an increase in n and this superior performance of stPCA also holds, if the spatio-temporal factor consists of a spatial trend instead of a patch.

As stated before, any spatial principal component approach will also try to identify local structure and report this structure as a negative eigenvalue. In the current simulation, which does not explicitly include local structure, this feature appears twice. At first, stPCA and the purely spatial PCA variants perform worse than the original PCA on non-spatial or only slightly spatial data, as can be observed by ratio which correspond to $\delta = 0$. Secondly this search for local structure causes

stPCA to possess a pronounced standard deviation as depicted in Figure 1, which however disappears as the spatial increment grows.

In a second simulation we apply PCA, its spatial variants and stPCA to a data set with two different spatio-temporal factors in order to learn the accuracy of these principal component approaches. Hence we extend the preceding simulation by an additional spatio-temporal factor, which is based on the spatial structure $\mathbf{S}^{(2)}$:

$$\begin{aligned} S_{i,r}^{(2)} &\sim \mathcal{N}(0, 1), \text{ for } r \leq \frac{R}{2} \\ S_{i,r}^{(2)} &\sim \mathcal{N}(\delta, 1), \text{ for } r > \frac{R}{2}, \end{aligned}$$

where δ describes the spatial increment, r denotes a row indicator and $R = \sqrt{n}$ indicates the number of rows. Consequently the spatial structure $\mathbf{S}^{(2)}$ describes a spatial patch, which differentiates between the upper and lower half of the grid.

The resulting structure defines the spatial distribution of the second spatio-temporal factor $F_{i,t}^{(2)}$ by serving as a constant in the respective AR(1) process:

$$\mathbf{F}_t^{(2)} = \mathbf{S}^{(2)} + 0.5\mathbf{F}_{t-1}^{(2)} + \boldsymbol{\varepsilon}_t,$$

where $\boldsymbol{\varepsilon}_t$ is defined as above. As in the preceding simulation, the hidden spatio-temporal factor $\mathbf{F}_t^{(2)}$ is observed via three noisy variables, which differ, as before, in their error components. In this context we set the white noise of variables affected by the first factor to $\boldsymbol{\varepsilon}_{\mathbf{Z},t}^{(1)} \sim \mathcal{N}(\mathbf{0}_n, 0.375\mathbf{I}_n)$ and the error vector of the variables defined by the second factor to $\boldsymbol{\varepsilon}_{\mathbf{Z},t}^{(2)} \sim \mathcal{N}(\mathbf{0}_n, 1.125\mathbf{I}_n)$.

Furthermore we add three additional iid noise processes and consequently observe three variables affected by the spatio-temporal factor $\mathbf{F}_t^{(1)}$, three variables influenced by spatio-temporal factor $\mathbf{F}_t^{(2)}$ and six additional noise variables, which complicate the disclosure of the two spatio-temporal factors.

As before we run this simulation in two settings to account for small ($n_1 = 49$) and large ($n_2 = 400$) data sets and allow for varying time dimensions: $t_1 \in \{5, 50, 100\}$ and $t_2 \in \{40, 400, 800\}$. We increase the spatial increment gradually in the interval $[0, 8]$ to observe its effect, make use of the aforementioned spatial weight matrix, and run each combination of parameters 1000 times.

We begin our inspection of the simulation results by assessing the power of the diverse principal component approaches to identify the correct number of factors. Figure 2 and Figure 3 presents modified scree plots for PCA, sPCA and stPCA in which the mean eigenvalues are interpolated to allow for several spatial increments δ to be shown in the same graph.

At first we observe that the distinct principal component approaches return different eigenvalues. The original PCA does not react to an increase in δ , as this non-spatial approach presents nearly the same eigenvalues for all levels of δ . On the other hand sPCA and stPCA do respond to an increase in the spatial increment. Especially the two largest and the smallest eigenvalues increase with δ and their reaction is amplified by more observations.

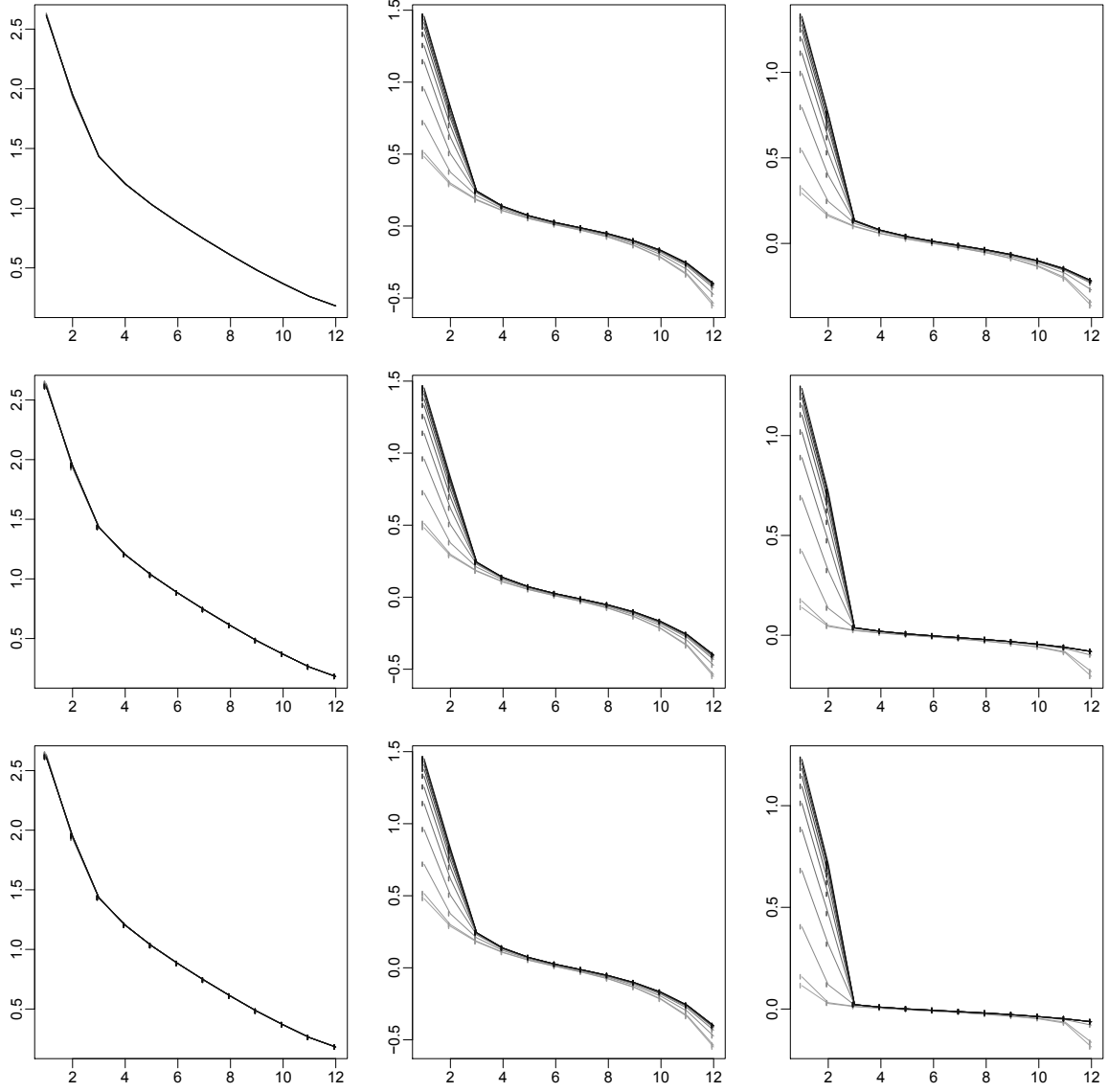


Figure 2: *Scree plots for PCA (first column), sPCA (second column) and stPCA (third column) depicted for $n_1 = 49$, the time frame $t_1 = 5$ (first row), $t_1 = 50$ (second row), $t_1 = 100$ (third row) and the spatial increments $\delta \in [0, 8]$ indicated by an increase in the grey strength.*

At $n_1 = 49$ the scree plot of PCA presents a slight decrease in the gradient starting at the third eigenvalue and consequently indicates the presence of two factors. These factors arise due to the constant in the simulation's AR(1) process. At $n_2 = 400$ PCA presents three obvious changes in the gradient resulting in two or more factors and consequently the scree plot does not clearly indicate the correct number of factors.

In contrast sPCA and stPCA react explicitly to an increase in δ . At $\delta \leq 0.8$ the eigenvalues returned by sPCA do not indicate any obvious change in the gradient, but rather suggest a smooth curve. Only at $\delta > 0.8$ one can make out a clear change in the curvature after the second eigenvalue and consequently sPCA identifies the two spatio-temporal factors. This finding is even more apparent as δ and n are increased. However, sPCA also returns large negative eigenvalues, which

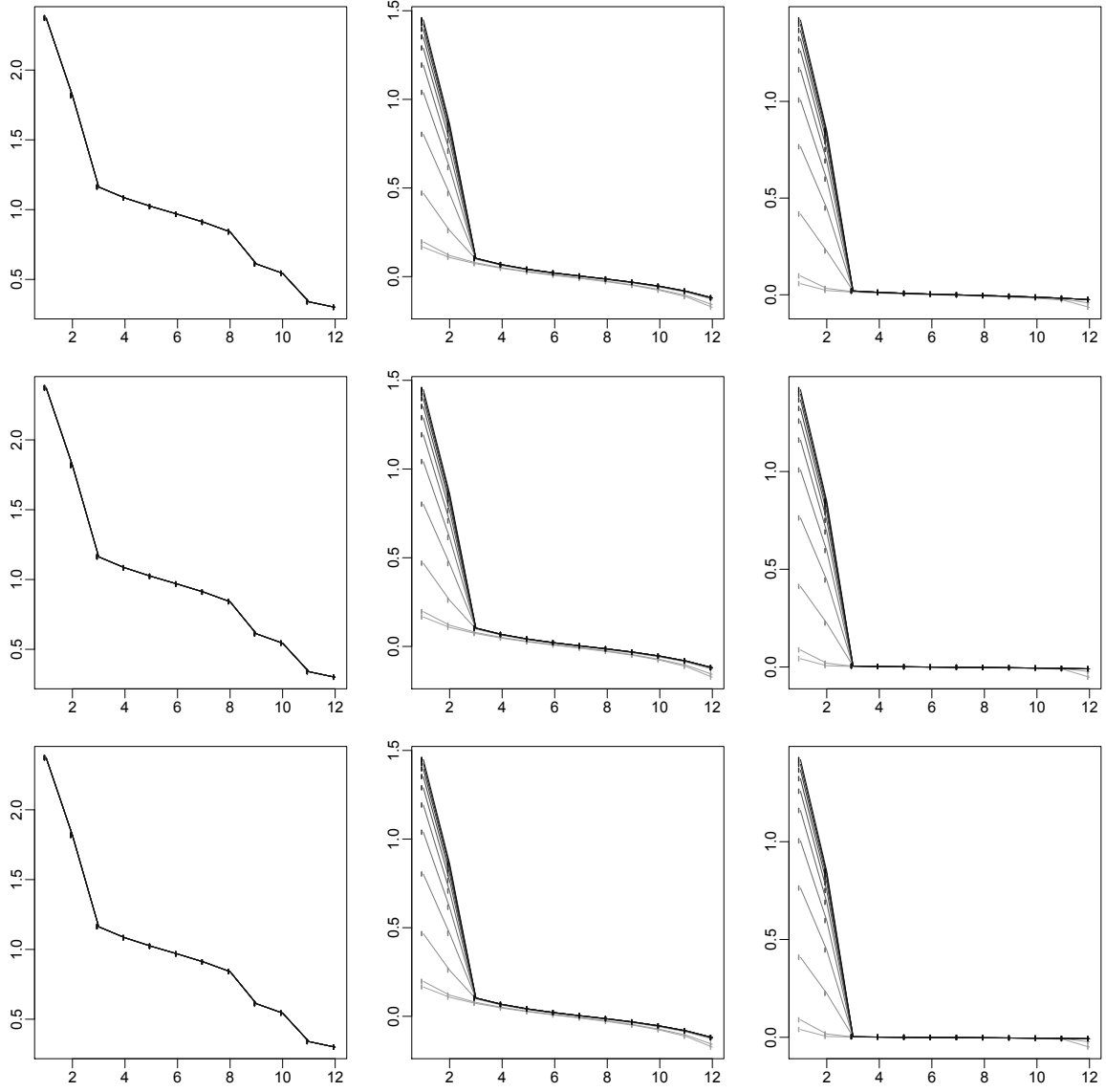


Figure 3: *Scree plots for PCA (first column), sPCA (second column) and stPCA (third column) depicted for $n_1 = 400$, the time frame $t_1 = 40$ (first row), $t_1 = 400$ (second row), $t_1 = 800$ (third row) and the spatial increments $\delta \in [0, 8]$ indicated by an increase in the grey strength.*

erroneously indicate a high level of local structure.

The application of stPCA results in eigenvalues, which possess a similar structure as sPCA, but which is much more pronounced. At the low level of $\delta \leq 0.8$ the respective eigenvalues also suggest a smooth curve without any clear indication of the number of factors. At $\delta > 0.8$ stPCA increasingly indicates the presence of the two factors. However the difference between the second and third eigenvalue is much wider in the case of stPCA than sPCA, e.g. the ratio at $\delta = 4$, $n_2 = 400$ and $t_2 = 400$ for stPCA (1:151.071) surpasses the ratio of sPCA (1:7.163) more than twentyfold and consequently stPCA indicates the two factors more evidently than sPCA. Furthermore stPCA indicates the presence of local structure only at very low levels of δ and the clarity of its scree plot is not only amplified by an increase in the spatial increment δ and the sample size n , but also by the time dimension t .

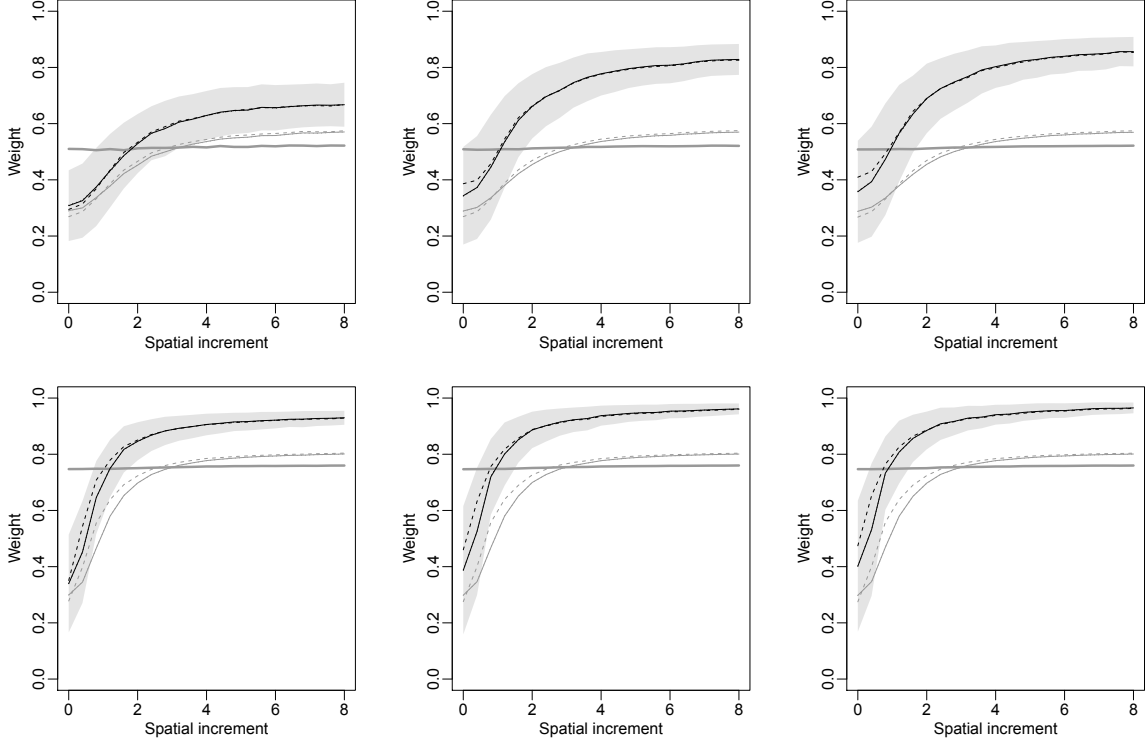


Figure 4: Mean weight (with standard deviation for temporal sPCA) assigned by PCA (thick solid grey line), sPCA (solid grey line), MSC (dashed grey line), Global Structure PCA (dotted grey line), temporal sPCA (solid black line), temporal MSC (dashed black line) and temporal Global Structure PCA (dotted black line) in the two eigenvectors to the respective variables for $n_1 = 49$ (first row) with $t_1 = 5$ (left graph), $t_1 = 50$ (middle graph) and $t_1 = 100$ (right graph), and $n_2 = 400$ (second row) with $t_2 = 40$ (left graph), $t_2 = 400$ (middle graph) and $t_2 = 800$ (right graph).

The scree plots of the spatial or spatio-temporal variants of MSC or Global Structure present a very similar pattern in the respective eigenvalues and we consequently do not present them here.

In a second evaluation step, we examine the eigenvectors which correspond to the two largest eigenvalues. These specify the weights in the linear combination of the variables to obtain the projections. In our simulated data the first spatio-temporal factor affects only the first three variables, whereas the second spatio-temporal factor defines the fourth, fifth and sixth variables. Consequently we would expect the first eigenvector to carry large weights on the first, second and third variable and the second eigenvector to accentuate the fourth, fifth and sixth variable. Furthermore, an optimal principal component approach would assign zero weights to all the other variables, as these do not contain any information on the two spatio-temporal factors.

Figure 4 presents the mean weight the two eigenvectors assign to the respective variables, that is the mean of the weight assigned by the first eigenvector to the first, second and third variable and the weight given by the second eigenvector to the fourth, fifth and sixth variable. Obviously a high ratio indicates that the respective principal component approach does not erroneously highlight variables, which do

not incorporate any information on the spatio-temporal factors.

At first we observe, that the original PCA attributes substantial weights to the respective variables, and the weights increase with n . However, PCA does not react to an increase in either the spatial increment δ nor the time dimension t and produces constant weights in this respect. Contrary to this indifference the spatial variants do react to an increase in δ and the corresponding weights surpass the original PCA's weights at $\delta \geq 3.2$ for $n_1 = 50$, respectively $\delta \geq 2.8$ for $n_2 = 400$. A further increase in the spatial increment widens the margin between PCA and the spatial PCA variants even more. But the spatial PCAs do not react to an increase in t and due to their cross-sectional approach they fail to exploit the serial correlation in order to improve the weights even further.

This is instead accomplished by the new stPCA procedure. The mean weights presented by this spatio-temporal technique improve upon an increase of either the spatial increment δ , the number of observations n and also upon an increase in the time dimension t . stPCA exceeds the weights of the alternative approaches at a spatial increment between $\delta \geq 2$ ($n_1 = 49$ and $t_1 = 5$) and $\delta \geq 0.8$ ($n_2 = 400$ and $t_2 = 800$), and this difference is amplified as t is increased. For example, at a parametrization of $n_2 = 400$, $t_2 = 400$ and $\delta = 4$ the mean weight of the eigenvectors presented by stPCA exceeds the weights of second-best procedure by 19.5%.

In a third evaluation step we directly compare the spatio-temporal factors with its projections. In detail we compute the mutual information $\text{MI}(\cdot, \cdot)$ between the artificially created factor values $\mathbf{F}_t^{(1,2)}$ and the projections $\hat{\mathbf{F}}_t^{(1,2)}$ identified by the diverse principal component approaches. The computation is based on the two largest positive eigenvalues of every principal component approach. Figure 5 reports on the average mutual information over time, which we define as

$$\frac{1}{2T} \sum_{t=1}^T \max \left\{ \text{MI} \left(\hat{\mathbf{F}}_t^{(1)}, \mathbf{F}_t^{(1)} \right) + \text{MI} \left(\hat{\mathbf{F}}_t^{(2)}, \mathbf{F}_t^{(2)} \right), \text{MI} \left(\hat{\mathbf{F}}_t^{(1)}, \mathbf{F}_t^{(2)} \right) + \text{MI} \left(\hat{\mathbf{F}}_t^{(2)}, \mathbf{F}_t^{(1)} \right) \right\}$$

to address issues arising from factor switching.

It might first be noted from Figure 5, that all PCA approaches struggle with a more and more pronounced spatial patch. All techniques report a decrease in the mutual information, if the spatial increment δ is increased up to its maximum. This observation can be accredited to the particular simulation setting in which the spatial patch is obscured by normally distributed noise. Consequently this error component prohibits a clearer identification of the increasing spatial structure, as the diverse PCA procedures weight the key variables and hence their accompanying noise stronger, as δ is increased. However, this effect is not observed, if a spatial trend is modelled instead of a patch and might ultimately be explained by the particular spatial structure.

The original PCA approach explains nearly half of the entropy of the factor values. It performs only marginally better, if the number of observation n is increased and does not react to the time dimension. As explained above its performance worsens as the spatial increment δ is raised.

The spatial variants of PCA surpass its ordinary cousin at $\delta \geq 2.8$, but its performance depends on the level of the spatial increment. An increase in δ at low

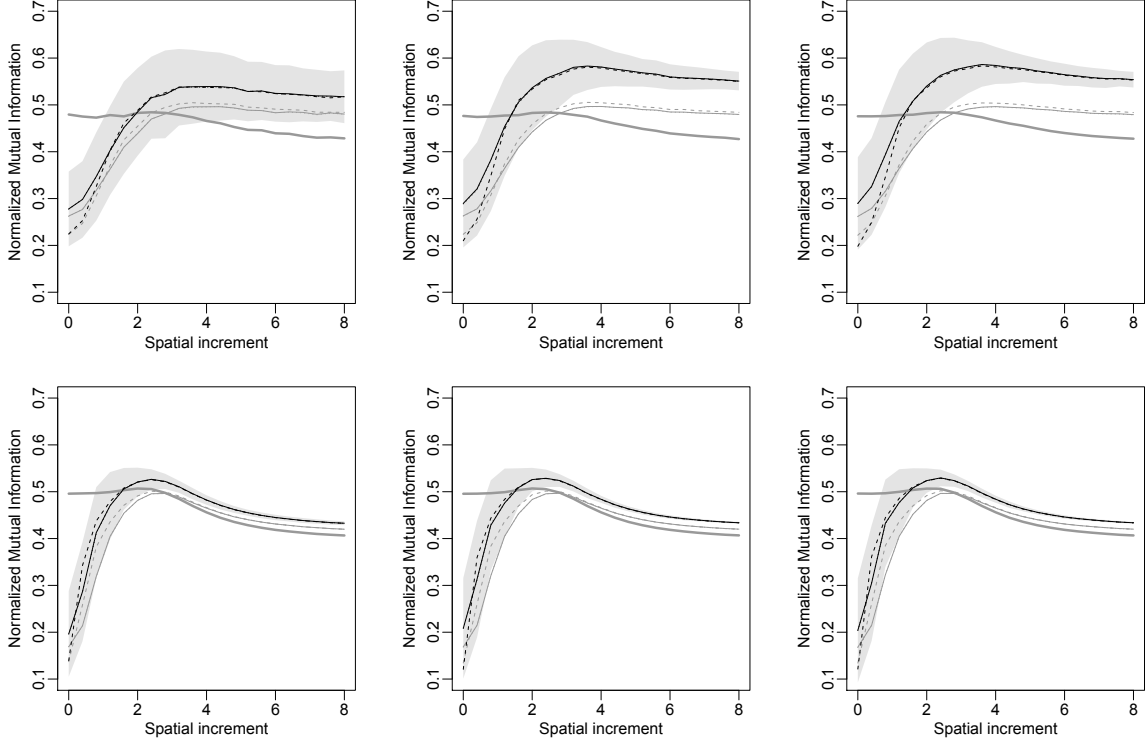


Figure 5: Normalized mutual information (with standard deviation for temporal sPCA) between the simulated factor values and the factor scores indicated by PCA (thick solid grey line), sPCA (solid grey line), MSC (dashed grey line), Global Structure PCA (dotted grey line), temporal sPCA (solid black line), temporal MSC (dashed black line) and temporal Global Structure PCA (dotted black line) for $n_1 = 49$ (first row) with $t_1 = 5$ (left graph), $t_1 = 50$ (middle graph) and $t_1 = 100$ (right graph), and $n_2 = 400$ (second row) with $t_2 = 40$ (left graph), $t_2 = 400$ (middle graph) and $t_2 = 800$ (right graph).

levels will strengthen its performance up to a maximum and a further increase in δ will afterwards worsen the mutual information between the factor values and its projections. The same observation holds for an increase in the number of observations n and as before the spatial PCA approaches do not respond to an increase in t .

In contrast to the ordinary and spatial PCA variants stPCA exploits an increase in the time dimension and the corresponding mutual information exceeds the alternative ones at $\delta \geq 2$ ($n_1 = 49$), respectively $\delta \geq 1.6$ ($n_2 = 400$). However like the spatial PCA approaches its performance on an increase in δ and n is mixed. On low levels of the spatial increments the mutual information rises steeply, if either δ or n are increased, but on high levels of δ a further increase of either the spatial increment or the number of observation worsen the mutual information of stPCA.

However, stPCA outperforms the ordinary and spatial PCA approaches on all parameter values apart from very low levels of δ and the difference is significant. Obviously the largest gain is made on small n , high t samples, whereas the additional value for large n data seems less pronounced.

In order to visualize the observed difference in the mutual information, Figure 6 presents the simulated factor values and corresponding projections returned by the distinct principal component approaches for $n_2 = 400$, $t_2 = 40$ and $\delta = 4$ at an

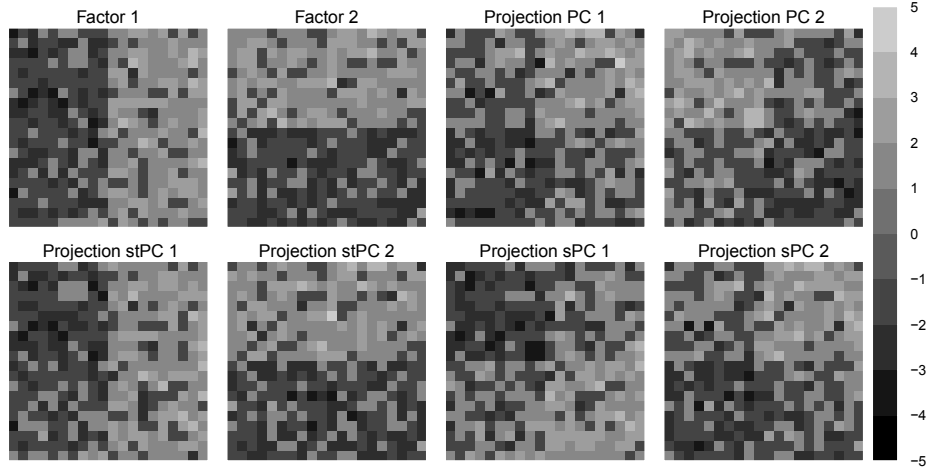


Figure 6: *Simulated factor values and projections by the respective principal component approaches for $n_2 = 400$, $t_2 = 40$ and $\delta = 4$ for an exemplary point in time.*

exemplary point in time. In order to get a better contrast in the graph, we added, respectively subtracted, 1 from the original values and projections.

The original factor values present the aforementioned patch between either the left and right or upper and lower half of the grid. Whereas the single projections by stPCA do not resemble the original factors perfectly, as a whole they clearly present the same basic structure of two diverse spatial structures. This observation does not hold for either the original PCA nor its spatial variants, as both techniques present a spatial structure which clearly discriminates between four disjunct parts in every corner, but fail to present the original spatial structure of one north-south and one west-east patch. Consequently and as noted above both techniques incorporate lower mutual information values.

4 Application on urbanism and economic deprivation

Apart from case-by-case specific impact factors, criminological theory and research based on data for areal units have persistently and mainly in the United States identified two broad dimensions of social structure that have proven to be robust predictors of violent crime rates: (1) economic well-being/relative deprivation, (2) population structure/urbanism (McCall et al., 2010). In a recent study, Messner et al. (2013) have confirmed the explanatory power of these two factors with regard to German assault and robbery rates collected in 413 geographic-administrative districts called “Kreise” (counties) and aggregated over the years 2005-2007.

In this study, the factor Urbanism was constructed via PCA performed on four indicators: (1) the average population size across the three-year period; (2) population density, i.e., population per square kilometer; (3) the proportion of the workforce employed in agriculture or forestry; (4) percent of the foreign-born population, which tends to be concentrated in urban areas.

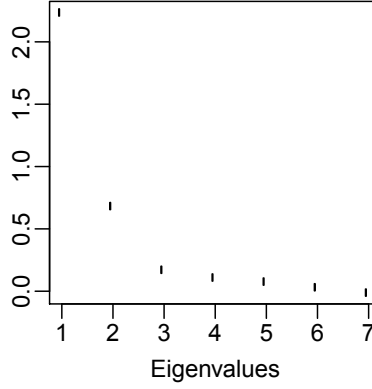


Figure 7: *Scree plot for the indicators on urbanism and economic deprivation.*

The factor economic well-being/deprivation was constructed via PCA performed on five indicators: (1) the percentage of the civilian labor force that is unemployed; (2) the percentage of those persons who receive social assistance; (3) average monthly household income; (4) the proportion of households receiving housing assistance; (5) a measure of high school dropouts.

In our own analysis we have removed the fifth indicator from the economic deprivation factor, since in a subsequent study using roughly the same data base Thome and Stahlschmidt (2013) have presented theoretical arguments and empirical evidence demonstrating that high school attendance should be included, together with three additional indicators, in another factor labelled “disintegrative individualism.” Since our present article does not deal with theoretical issues we have limited our analysis, i.e. the application of stPCA, to the two factors most commonly applied in criminological research based on areal units, economic deprivation and urbanism. Furthermore we exclude the percentage of foreign-born population, as this variable is closely related to both factors, urbanism and relative deprivation, and therefore interferes with our aim to construct two clearly distinguished factors.

Whereas Messner et al. (2013) averaged their data over a three-year period to generate the factor scores in a cross-sectional approach, stPCA allows to include every year separately in the analysis. Hence, the presented projections are based on more data points and take into account serial correlation.

Figure 7 presents the scree plot generated by stPCA and based on a row weighted binary spatial weight matrix indicating direct neighbouring counties. An aggregation of the seven indicators into two factors seems reasonable, as the first two eigenvalues clearly stand out, if compared to the remaining ones.

The corresponding variable weights for each factor are detailed in Table 1. The variables are grouped as expected by criminological theory. The unemployment rate, social welfare rate, household income and housing allowance rate together describe a factor interpreted as relative economic deprivation, whereas population size, population density and employment in agriculture jointly specify the level of urbanism. Most of the single variables can clearly be attributed to one of the two factors. Only the social welfare rate takes on similar weights in both factors since the need for

	Eco. Deprivation	Urbanism
Unemployment rate	0.563	0.222
Social welfare rate	0.446	0.339
Household income	-0.465	0.293
Housing allowance rate	0.503	-0.092
Population size	-0.013	0.396
Population density	-0.060	0.554
Employment in agriculture	0.105	-0.527

Table 1: *The first and second eigenvector corresponding to the two largest eigenvalues.*

such payments arises less often in rural areas.

Finally Figures 8 and 9 display the projections resulting from the application of the weights to the single counties. The projections of the economic deprivation factor depict two spatial patterns. First, there remains a clear disparity between East and West Germany, as counties of the former German Democratic Republic possess less economic means than their western counterparts. Second, inside these two blocks there arises a North–South divide as the southern part of Germany has, during the last two decades, achieved a more advanced, higher-level balance between traditional and modernized sectors of economic development. These two patterns remain stable over the inspected time horizon. This stationarity over time is also observed in case of the urbanism factor. The corresponding projections clearly highlight the big urban hubs of Berlin, Munich, Hamburg or the Rhine-Ruhr metropolitan region and the sparsely populated north-eastern part of Germany.

5 Conclusion

The analyses of purely spatial data are confronted with the implied peculiarities of such data, namely spatial correlation and heterogeneity. Consequently such analyses entail a need for a high amount of data in order to obtain reliable estimates of any parameter of interest. But due to natural limits, which are especially obvious in the case of areal data, sample sizes can not be enlarged indefinitely over space. However, a feasible solution in dealing with this problem can be forged ahead by the extension of such data over time resulting in spatio-temporal analyses. This approach requires the transformation of well-known instruments and techniques from the i.i.d. or spatial environment to the spatio-temporal one.

To our knowledge, stPCA offers a first attempt to transfer the original PCA to the spatio-temporal realm of geographical and serial correlation. The proposed technique allows for dimension reduction on the attribute space while preserving the geographical and temporal structure and it offers a promising approach to generate consistent factors from spatio-temporal data. It differs from any explicit factor modelling approach by its algorithmic nature, which can be viewed as a welcomed feature or a drawback depending on substantive issues given in a particular research context.

In any case stPCA possesses a superior performance in terms of sensibility to detect

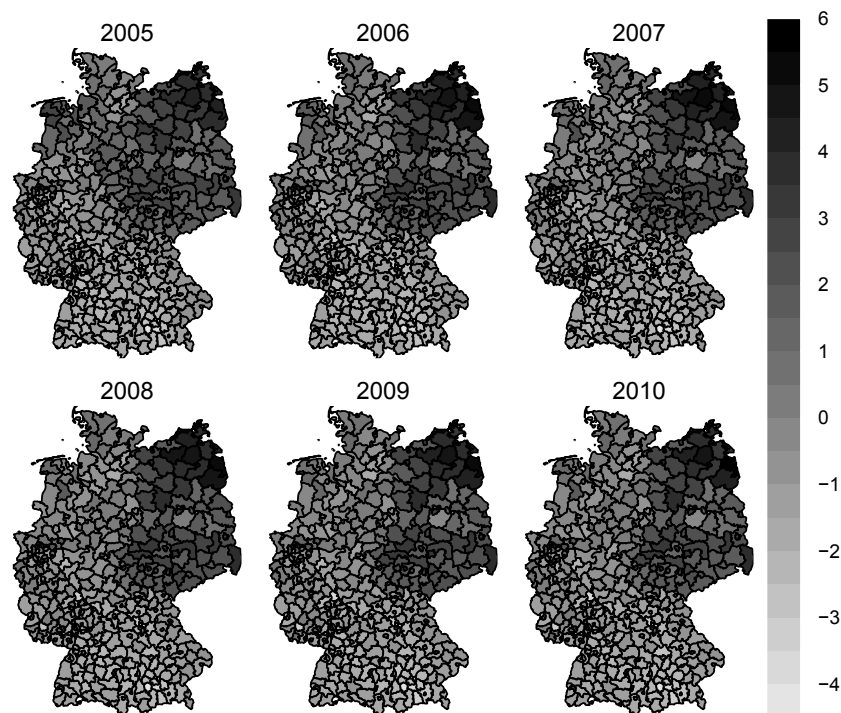


Figure 8: *Projections of the factor “Economic Deprivation” on German Kreise. Digital map provided and copyrighted by GeoBasis-DE / BKG 2013*

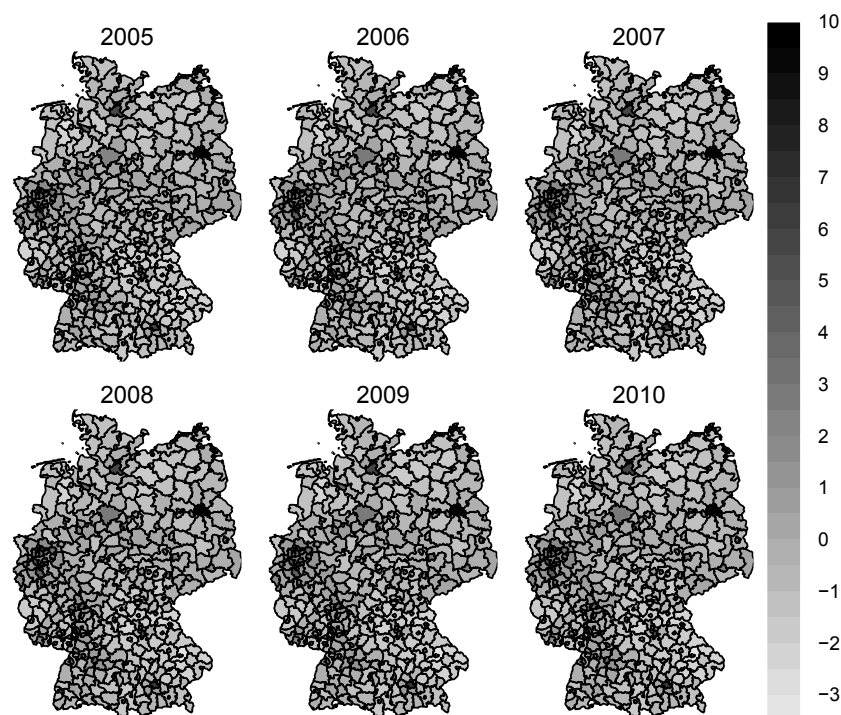


Figure 9: *Projections of the factor “Urbanism” on German Kreise. Digital Map provided and copyrighted by GeoBasis-DE / BKG 2013*

and of accuracy to disclose spatio-temporal factors, if compared to the original PCA and the proposed spatial variants thereof. Especially the original PCA lacks power to correctly identify spatial factors and its spatial variants fail to exploit any serial correlation to improve their results due to their static nature. Furthermore stPCA is much faster than its archetypes, as the time-consuming eigendecomposition has to be calculated only once instead of t times.

As PCA and the spatial variants have to be applied separately for every t , they are prone to sign and factor switching over t . This behaviour complicates an analysis, as such instances have to be detected and resolved before the projections can be analysed or supplied for further use. stPCA avoids such issues, as it presents time stable weights for the linear combinations and consequently allows for a direct and consistent interpretation of these values. Summing up, stPCA seems to be better suited than its static forerunners to address the specific requirements arising from spatio-temporal analyses.

However, the projections resulting from stPCA obviously depend on the variable scale and on the appropriateness of the spatial weight matrix. Furthermore the performance of stPCA might be amplified by a suitable orthogonal or oblique rotation, which will consequently restrict the algorithmic nature of stPCA and increase the researcher's influence.

Finally, we would like to mention two modifications of stPCA, from which certain applications might benefit. First, the spatial weight matrix could be understood as time dependent and, upon availability, a time-specific W_t could be introduced into the optimisation (3) to refine the technique. Second, stPCA could also be developed into an adaptive approach, in which the projections for t are not obtained via the time average of the spatial covariance matrix over the whole time frame T , but only over the interval $[t - t^*, t + t^*]$, where $t^* < T/2$ denotes a tuning parameter. An appropriate weighting schema over this interval could furthermore enhance the flexibility. This moving window approach resides between the spatial variants of PCA and stPCA, as it exploits serial correlation over time, but foregoes any computational advantage, as an eigendecomposition has to be obtained at every t separately. Yet this adaptive stPCA might allow for time trends in the data and consequently account for not only the spatial peculiarities in spatial-temporal data, but also for the temporal ones.

References

- Choi, J., Lawson, A.B., Cai, B., Hossain, M.M., Kirby, R.S. and Liu, J. (2012). A Bayesian latent model with spatio-temporally varying coefficients in low birth weight incidence data. *Statistical Methods in Medical Research*, **21**, 445–456.
- Fotheringham, A.S., Brunsdon, C. and Charlton, M. (2002). *Geographically Weighted Regression: The Analysis of Spatially Varying Relationships*. Chichester: Wiley.
- Gelman, An and Hill, J. (2006). *Data Analysis Using Regression and Multi-level/Hierarchical Models*. Cambridge: Cambridge University Press.
- Härdle, W.K. and Simar, L. (2012). *Applied Multivariate Statistical Analysis* (3rd ed.). Berlin: Springer.
- Hogan, J.W. and Tchernis, R. (2004). Bayesian Factor Analysis for Spatially Correlated Data, With Application to Summarizing Area-Level Material Deprivation From Census Data. *Journal of the American Statistical Association*, **99**, 314–324.
- Hotelling, H. (1933). Analysis of a Complex of Statistical Variables into Principal Components. *Journal of Educational Psychology*, **24**, 417–441 & 498–520.
- Jolliffe, I.T. (2002). *Principal Component Analysis*. New York: Springer.
- Jombart, T., Devillard, S., Dufour, A.-B. and Pontier, D. (2008). Revealing cryptic spatial patterns in genetic variability by a new multivariate method. *Heredity*, **101**, 92–103.
- Karhunen, K. (1947). Über lineare Methoden in der Wahrscheinlichkeitsrechnung. *Ann. Acad. Sci. Fennicae, ser. A1, Math. Phys.*, **37**.
- Kroonenberg, P.M. and de Leeuw, J. (1980). Principal Component Analysis of Three-Mode Data by Means of Alternating Least Squares Algorithms. *Psychometrika*, **45**, 69–97.
- Lawson, A.B., Song, H.-R., Cai, B., Hossain, M.M. and Huang, K. (2008). Space-time latent component modeling of geo-referenced health data. *Statistics in Medicine*, **29**, 2012–2027.
- Loève, M. (1948). Fonctions Aléatoires de second order. In Lévy, P. (ed.), *Processus Stochastique et Movement Brownien*, Paris: Hermann.
- Lorenz, E. (1956). Empirical orthogonal functions and statistical weather prediction. *Statistical Forecasting Project*, Scientific Report 1.
- McCall, P.L., Land, K.C., and Parker, K.F. (2010). An empirical assessment of what we know about structural covariates of homicide rates: A return to a classic 20 years later. *Homicide Studies*, **14**, 219–243.
- Messner, S.F., Teske, R.H.C., Baller, R.D., and Thome, H. (2013). Structural covariates of violent crime rates in Germany: Exploratory spatial analyses of Kreise. *Justice Quarterly*, **30**, 1015–1041.
- Moran, P.A.P. (1950). Notes on Continuous Stochastic Phenomena. *Biometrika*, **37**, 17–23.

- Pearson, K. (1901). On Lines and Planes of Closest Fit to Systems of Points in Space. *Philosophical Magazine, Series 6*, **2**, 559–572.
- Richman, M.B. (1986). Rotation of principal components. *Journal of Climatology*, **6**, 195–216.
- Spearman, C. (1904). “General Intelligence,” objectively determined and measured. *American Journal of Psychology* **15**, 201–293.
- Thioulouse, J., Chessel, D. and Champely, S. (1995). Multivariate Analysis of spatial patterns: a unified approach to local and global structures. *Environmental and Ecological Studies*, **2**, 1–14.
- Thome, H. and Stahlschmidt, S. (2013). Ost und West, Nord und Süd: Zur räumlichen Verteilung und theoretischen Erklärung der Gewaltkriminalität in Deutschland. *Berliner Journal für Soziologie*, **23**, 441–470.
- Tzala, E. and Best, N. (2007). Bayesian latent variable modelling of multivariate spatio-temporal variation in cancer mortality *Statistical Methods in Medical Research*, **17**, 97–118.
- Wartenberg, D. (1985). Multivariate Spatial Correlation: A Method for Explanatory Geographical Analysis. *Geographical Analysis*, **17**, 263–282.

SFB 649 Discussion Paper Series 2014

For a complete list of Discussion Papers published by the SFB 649, please visit <http://sfb649.wiwi.hu-berlin.de>.

- 001 "Principal Component Analysis in an Asymmetric Norm" by Ngoc Mai Tran, Maria Osipenko and Wolfgang Karl Härdle, January 2014.
- 002 "A Simultaneous Confidence Corridor for Varying Coefficient Regression with Sparse Functional Data" by Lijie Gu, Li Wang, Wolfgang Karl Härdle and Lijian Yang, January 2014.
- 003 "An Extended Single Index Model with Missing Response at Random" by Qihua Wang, Tao Zhang, Wolfgang Karl Härdle, January 2014.
- 004 "Structural Vector Autoregressive Analysis in a Data Rich Environment: A Survey" by Helmut Lütkepohl, January 2014.
- 005 "Functional stable limit theorems for efficient spectral cointegration estimators" by Randolph Altmeyer and Markus Bibinger, January 2014.
- 006 "A consistent two-factor model for pricing temperature derivatives" by Andreas Groll, Brenda López-Cabrera and Thilo Meyer-Brandis, January 2014.
- 007 "Confidence Bands for Impulse Responses: Bonferroni versus Wald" by Helmut Lütkepohl, Anna Staszewska-Bystrova and Peter Winker, January 2014.
- 008 "Simultaneous Confidence Corridors and Variable Selection for Generalized Additive Models" by Shuzhuan Zheng, Rong Liu, Lijian Yang and Wolfgang Karl Härdle, January 2014.
- 009 "Structural Vector Autoregressions: Checking Identifying Long-run Restrictions via Heteroskedasticity" by Helmut Lütkepohl and Anton Velinov, January 2014.
- 010 "Efficient Iterative Maximum Likelihood Estimation of High-Parameterized Time Series Models" by Nikolaus Hautsch, Ostap Okhrin and Alexander Ristig, January 2014.
- 011 "Fiscal Devaluation in a Monetary Union" by Philipp Engler, Giovanni Ganelli, Juha Tervala and Simon Voigts, January 2014.
- 012 "Nonparametric Estimates for Conditional Quantiles of Time Series" by Jürgen Franke, Peter Mwita and Weining Wang, January 2014.
- 013 "Product Market Deregulation and Employment Outcomes: Evidence from the German Retail Sector" by Charlotte Senftleben-König, January 2014.
- 014 "Estimation procedures for exchangeable Marshall copulas with hydrological application" by Fabrizio Durante and Ostap Okhrin, January 2014.
- 015 "Ladislaus von Bortkiewicz - statistician, economist, and a European intellectual" by Wolfgang Karl Härdle and Annette B. Vogt, February 2014.
- 016 "An Application of Principal Component Analysis on Multivariate Time-Stationary Spatio-Temporal Data" by Stephan Stahlschmidt, Wolfgang Karl Härdle and Helmut Thome, February 2014.

SFB 649, Spandauer Straße 1, D-10178 Berlin
<http://sfb649.wiwi.hu-berlin.de>

This research was supported by the Deutsche
Forschungsgemeinschaft through the SFB 649 "Economic Risk".

